

RHIC & USAtlas Computing Facilities Needs for Development

The RHIC and USAtlas Computing Facilities (RACF) at Brookhaven Nat Lab are DOE user facilities providing computing resources to RHIC users and serving as national data storage and data processing center for the US based participants in the LHC Atlas program.

The Relativistic Heavy Ion Collider (RHIC) – a world-class nuclear physics research facility aimed at studies of nuclear matter at high energy and density -- began operations in 2000 and already collected more than 2PB of experimental data. The RHIC Computing Facility has more than 1000 users from 20 countries and delivered millions of hours of CPU in support of data reconstruction and physics analysis. However, the requirements of the user community continue to increase, demanding novel approaches to high performance data analysis that provide fast turnaround and near real time response, while geographically, the distribution of collaborating institutions continues to expand.

USAtlas pushes further the need to offer large volume data storage, high bandwidth networks and high performance computing to widely distributed collaborations. [...]

RACF operates in a tightly secure environment but has to provide global data access and high-bandwidth data channels to transport the datasets between the participating sites.

Our users and administrators need tools to:

- Access securely and easily systems of diverse nature: web services, login, data transfer services, personal communication and databases.
- Audit user activity and detect unacceptable usage.
- Provide accounting information for billing and resource allocation policy enforcement.
- Filter and protect high-bandwidth channels to allow massive transfers of data between sites without compromising site security.

The new collaborations require large scale data storage facilities at modest cost. The site storage systems should be interlinked to form a data storage grid offering integrated and transparent user service. The solution may evolve from the current Storage Resource Manager (SRM) developments providing low cost storage by integration with computing farm infrastructure, grid enabled interfaces and VO based accounting and allocation policies. The needs are:

- Future development of low cost storage managers based on shared computing farm infrastructure
- Development of grid enabled SRMs capable to enforce VO based allocation policies and provide detailed accounting information.
- Integration of network bandwidth management systems with storage managers to deliver predictable access to storage resources.

Due to a continuously increasing scale of the problems, most of our experimental groups need to use computing grids to analyze large volumes of data, located on a wide spread set of storage resources, without requiring every individual user have detailed knowledge of the data layout, CPU availability or grid access specifics. It is necessary to develop an integrated data analysis system, providing transparent access to resources and

empowering each researcher through access to the complete data inventory and computing capabilities.